



**VALEUR ALIMENTAIRE DES FOURRAGES LIGNEUX
CONSOMMÉS PAR LES RUMINANTS
EN AFRIQUE CENTRALE ET OCCIDENTALE**

RAPPORT FINAL

ALLEMAGNE	Universität Hohenheim - Institut für Tierernährung
BELGIQUE	CRA Gembloux - Station de Haute Belgique
CAMEROUN	IRZV - Yaoundé, Ngaoundéré, Garoua
CÔTE D'IVOIRE	IDESSA-DRA - Bouaké
FRANCE	CIRAD-EMVT - Maisons-Alfort CIRAD-FORÊT - Nogent/Marne INRA-SRNH - Theix
MALI	IER - Sotuba, Niono
SÉNÉGAL	ISRA-DRPSA-LNERV - Dakar
TCHAD	LRVZ - Farcha, N'Djaména
	et
BURKINA FASO	IDR - Ouagadougou CIRAD-FORÊT/IRBET - Ouagadougou
NOUVELLE-CALÉDONIE	CIRAD-EMVT - Nouméa

NOVEMBRE 1994

Centre de Coopération internationale en Recherche agronomique pour le Développement
Département d'Élevage et de Médecine vétérinaire
CIRAD-EMVT
10, rue Pierre-Curie 94704 Maisons-Alfort Cedex France



SOMMAIRE DU RAPPORT FINAL

Résumé

Chapitre I : Présentation succincte du programme d'étude de la valeur fourragère des arbres et arbustes d'Afrique tropicale centrale et occidentale

Chapitre II : Caractérisation des disponibilités fourragères ligneuses

Chapitre III : Composition botanique des régimes des ruminants sur parcours - Appétibilité relative des espèces ligneuses

Chapitre IV : Récolte et commercialisation des fourrages ligneux en régions périurbaines

Chapitre V : Echantillonnage des fourrages ligneux - Analyses au laboratoire - Composition chimique et dégradabilité enzymatique

Chapitre VI : Les tanins dans les fourrages ligneux

Chapitre VII : *In vitro* Untersuchungen zur Ermittlung der Verdaulichkeit, des Gehaltes an umsetzbarer Energie und des Proteinverfügbarkeit bei tropischen Futterbäumen und Büschen

In vitro studies for the prediction of digestibility, metabolisable energy content and protein fermentability of shrubs and tree fodders

(Mesures *in vitro* de la digestibilité pour estimer les teneurs en énergie métabolisable et en azote dégradable dans le rumen des fourrages ligneux)

Chapitre VIII : Prévion par la spectrométrie dans le proche infrarouge (SPIR) de la composition chimique et de la dégradabilité enzymatique ou *in vitro* des fourrages ligneux

Chapitre IX : Dégradabilité *in situ* de la matière sèche et des matières azotées de quelques fourrages ligneux : application des méthodes *in sacco* à l'étude de la digestion dans le rumen et dans les intestins

Chapitre X : Ingestion et digestibilité *in vivo* des fourrages ligneux

Chapitre XI : Essais d'alimentation de moutons avec des fourrages ligneux

Conclusion générale

Annexe I : Liste des espèces ligneuses enregistrées au catalogue CIRAD-EMVT/Laboratoire d'Ecologie végétale de l'Université Paris-Sud et des espèces ligneuses et subligneuses échantillonnées et étudiées sur l'animal

Human Resource Training - Equipment

→ **Liste des Abréviations**

CHAPITRE VIII

PREVISION PAR LA SPECTROMETRIE DANS LE PROCHE INFRAROUGE DE LA COMPOSITION CHIMIQUE ET DE LA DEGRADABILITE ENZYMATIQUE OU IN VITRO DES FOURRAGES LIGNEUX

Georges SINNAEVE*, Pierre DARDENNE*,
Philippe LECOMTE* et Robert BISTON*

* Station de Haute Belgique - Centre de Recherche Agronomique de Gembloux.
Ministère de l'Agriculture - rue de Serpont, 100 - B-6800 Libramont -
Belgique.

CCE.DGXII - ST2 - 1994. Prévision par la spectrométrie dans le proche infrarouge de la composition chimique et de la dégradabilité enzymatique ou *in vitro* des fourrages ligneux - Chapitre VIII In : GUERIN (H.) éd. Valeur alimentaire des fourrages ligneux consommés par les ruminants en Afrique centrale et de l'ouest. Commission des Communautés Européennes DG XII. Programme ST2/89/215. Maisons-Alfort, CIRAD-LEMVT.

SOMMAIRE

	Page
1. PRINCIPES ET METHODES	1
1.1. Aspects théoriques et fondamentaux	1
1.2. Principe de la mesure	1
1.3. Principe de l'étalonnage	2
1.4. Traitement des données spectrales	4
1.5. Détection d'échantillons aberrants	5
1.6. Les applications de la SPIR	5
1.7. Abréviations utilisées	6
2. APPLICATION AUX FOURRAGES LIGNEUX TROPICAUX : RESULTATS ET DISCUSSIONS	6
2.1. Etablissement de modèles généraux de prédiction : comparaison entre valeurs estimées et valeurs mesurées	8
2.1.1. Prédiction des composés organiques majeurs	8
2.1.2. Prédiction des composés azotés	12
2.1.3. Prédiction de la digestibilité <i>in vitro</i>	12
2.1.4. Prédiction des composés minéraux	19
2.2. Regroupement d'espèces sur base des données spectrales et analytiques : établissement d'équations spécifiques	20
2.2.1. Calibration sur base de groupes	20
2.2.2. Calibrations par famille et par genre	22
2.2.3. Sélections spectrales	23
2.3. Prédiction de la dégradabilité <i>in vitro</i> par le "gastest" (DMO_{GT}) à partir des valeurs de composition chimique ou de dégradabilité enzymatique prédites par SPIR	28
2.4. Essais sur une espèce importante : <i>Faidherbia albida</i> .	29
2.5. Essais d'un mode de calibration particulier la "Local Weighted Regression".....	29
3. ESSAIS SUR ANIMAUX : DEGRADABILITE <i>IN SACCO</i> , DIGESTIBILITE <i>IN VIVO</i>	32
CONCLUSION	33
REFERENCES BIBLIOGRAPHIQUES	35

1. PRINCIPES ET METHODES

Les objectifs qui ont été confiés à la Station de Haute Belgique dans le cadre de ce projet portent sur l'utilisation de la spectrométrie dans le proche infrarouge (SPIR) en vue de la détermination rapide des principaux constituants chimiques et de la valeur alimentaire des fourrages ligneux que l'on trouve dans les zones sahélo-soudaniennes de l'Afrique.

1.1. Aspects théoriques et fondamentaux

Les atomes d'une molécule forment des liaisons chimiques dont la fréquence de vibration est caractéristique d'un groupement atomique. Ces liaisons intramoléculaires sont susceptibles d'absorber tout rayonnement incident produit à la même fréquence de vibration.

La spectrométrie dans l'infrarouge proche est une méthode d'analyse instrumentale basée sur l'absorption par la matière d'une radiation électromagnétique dans le proche infrarouge (1 100 à 2 500 nm).

Dans cette gamme spectrale, les bandes d'absorption sont le fait de groupes atomiques tels que C-H, O-H, N-H. Les bandes en SPIR sont surtout des recombinaisons et des harmoniques des vibrations fondamentales s'opérant dans l'infrarouge moyen. La quantité de radiation lumineuse absorbée à chaque longueur d'onde suit la loi de Lambert-Beer et est proportionnelle au nombre de liaisons chimiques et donc à leur concentration. Après étalonnage, la SPIR peut être utilisée à des fins d'analyses quantitatives.

Certains constituants d'un même produit peuvent absorber la lumière dans l'infrarouge proche à une même longueur d'onde ou à des longueurs d'onde voisines. La SPIR ne peut être utilisée quantitativement que par traitements mathématiques des spectres. Les longueurs d'onde spécifiques des principaux constituants organiques sont reprises au tableau VIII.1 (Biston et Dardenne, 1985).

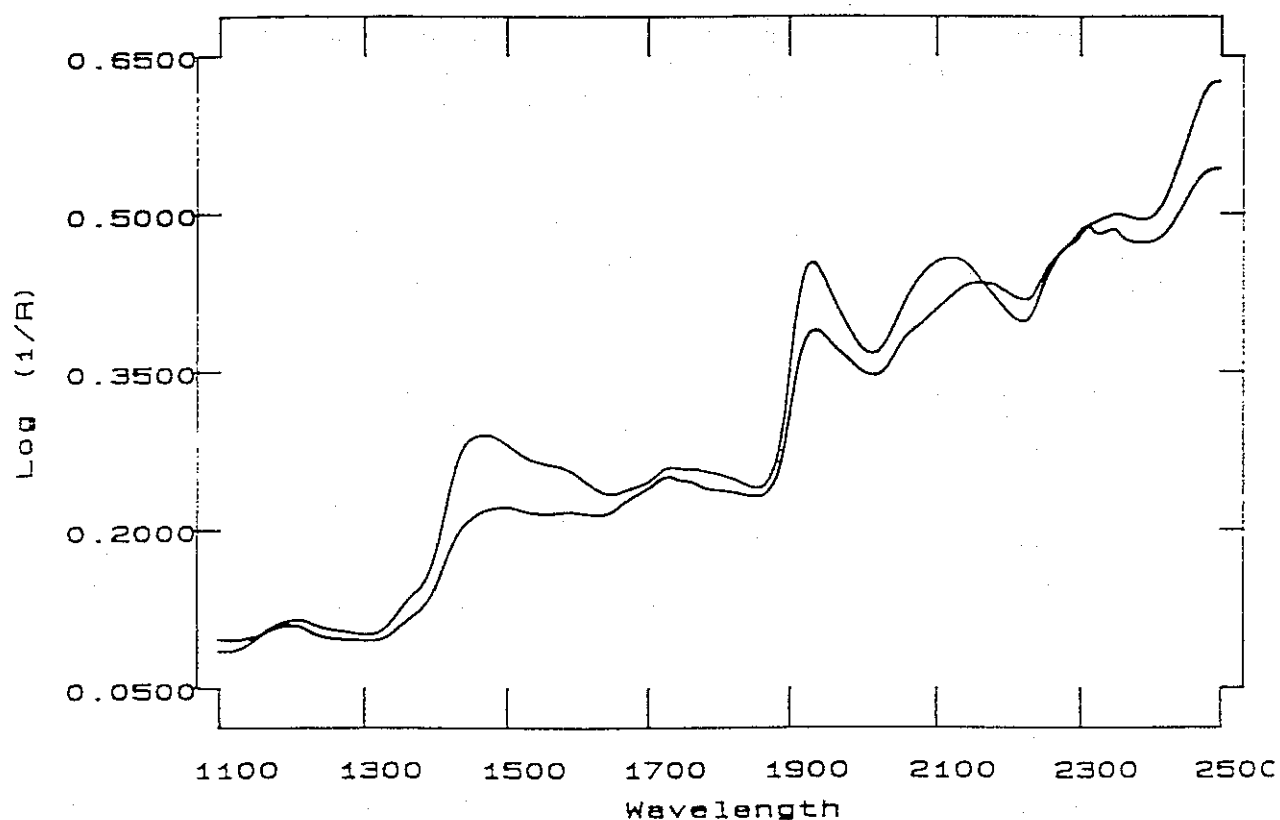
Tableau VIII.1 - Longueurs d'onde spécifiques des principaux constituants

Constituants	Longueurs
Eau	<u>1445</u> - 1790 - <u>1940</u>
Protéines	<u>1734</u> - 2060 - <u>2180</u> - 2300 - 2352 - 2470
Cellulose	<u>1722</u> - <u>1780</u> - <u>1818</u> - 1860 - 2100 - <u>2270</u> - <u>2335</u> - <u>2348</u> - 2488
Amidon	<u>1722</u> - <u>1780</u> - <u>2100</u> - <u>2270</u>
Matières grasses	<u>1759</u> - 2070 - <u>2140</u> - <u>2310</u> - 2380

1.2. Principe de la mesure

Après séchage et broyage standardisé (Cyclotec 1mm), l'échantillon est soumis au rayonnement proche infrarouge. La réflexion aux différentes longueurs d'onde constitue un spectre, véritable signature de l'échantillon dans la gamme de l'infrarouge. Cette "empreinte digitale" contient toute l'information physico-chimique de l'échantillon. Un exemple de spectres d'absorbance est repris à la figure VIII.1.

Figure VIII.1 - Spectres dans le proche infrarouge de fourrages tropicaux ligneux



1.3. Principe de l'étalonnage

Le dosage d'un constituant par SPIR nécessite une étape préalable d'étalonnage qui établit un modèle linéaire reliant les données spectrales à des dosages obtenus par méthode de référence.

Une base de calibration idéale est constituée d'échantillons représentatifs des échantillons dont on veut prédire la constitution. Cette base de données intégrera des facteurs tels que, conditions phytotechniques, années de récolte, lieux de récolte... En outre, les échantillons devront couvrir la plus large gamme de composition pour le constituant envisagé.

Deux procédures de calibration peuvent être appliquées à savoir le calibrage énumératif pour une population d'échantillons fermée et le calibrage analytique pour une population ouverte (tableau VIII.2.) (Dardenne, 1990).

Tableau VIII.2 - Comparaison entre calibrage énumératif et analytique
(Dardenne, 1990)

Procédure énumérative	Procédure analytique
1. Accumuler les échantillons	1. Obtenir les spectres des échantillons
2. Enregistrer les spectres de chacun des échantillons	2. Sélectionner un sous-ensemble d'échantillons intégrant le maximum de variabilité en vue de l'analyse laboratoire
3. Sélection aléatoire d'échantillons à analyse au laboratoire	3. Choisir les longueurs d'onde spécifiques du constituant à doser
4. Effectuer les calibrages sur les échantillons sélectionnés	4. Valider l'équation de calibrage par l'analyse d'échantillons de prédiction
5. Prédire la teneur des échantillons restants	5. Effectuer les analyses de routine

L'établissement du modèle prédictif reliant les densités optiques et les valeurs de référence est basé sur la recherche d'un modèle multilinéaire, qui exprimera le constituant en fonction des absorbances à certaines longueurs d'onde.

$$y = b_0 + b_1x_1 + b_2x_2 \dots + b_ix_i + \dots b_px_p$$

avec

- y = la concentration estimée
- b_0 = le terme constant
- b_i = les coefficients de régression
- x_i = les densités optiques ou leur transformée
- p = le nombre de termes

Dans cette équation, nous devons retrouver au moins une longueur d'onde caractéristique du constituant à doser. Les autres termes sont des termes correctifs intégrant les interférences de la granulométrie et des autres constituants.

La qualité d'une équation de calibrage est exprimée par l'écart-type de calibrage (SEC), l'écart-type de prédiction (SEP) (série indépendante) et le coefficient de corrélation multiple (R).

$$SEC = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-p-1}}$$

n = nombre d'échantillons
p = nombre de termes

$$SEP = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}$$

$$R = (1 - SCE_r/SCE_t)^{1/2}$$

SCE_r = somme des carrés des écarts résiduelle

SCE_t = somme des carrés des écarts totale

R indique l'étroitesse de la relation entre valeurs de référence et valeurs prédites. Son carré R^2 , ou coefficient de détermination, exprime la part de la variance observée de la variable dépendante qui est expliquée par la relation.

L'utilisation du rapport SD/SEC ou SD/SEP (SD = écart-type de la population) permet de comparer les performances des équations en faisant abstraction des unités.

R^2 , SEC et SD sont reliés entre eux par la relation suivante:

$$R^2 = 1 - \frac{SEC^2}{SD^2}$$

Le coefficient de détermination sera d'autant plus proche de 1 que SD sera grand et SEC petit, d'où l'importance du choix des échantillons pour obtenir une distribution rectangle sur les valeurs de référence.

R^2 = coefficient de détermination
SEC = erreur standard de calibrage
SD = écart-type de la population

1.4. Traitement des données spectrales

La reconnaissance de spectres intervient à deux niveaux. Dans un but de classement, les groupes étant distincts au départ, un nouvel individu est classé avec un minimum d'erreurs. Dans un but de classification au sein d'une matrice globale, on peut constituer des sous-groupes homogènes (Dardenne, 1990).

L'utilisation de l'analyse en composantes principales permet de réduire les spectres à un plus petit nombre de valeurs sans perte d'information. Une fois les facteurs principaux obtenus, les techniques telles que régression multiple, analyse discriminante et classification sont applicables sur les nouvelles coordonnées.

1.5. Détection d'échantillons aberrants

Lors de l'établissement de l'équation de calibrage un échantillon est considéré comme anormal lorsque la différence entre la teneur de référence et la valeur estimée est trop grande par rapport à l'écart-type résiduel. Ces échantillons sont écartés de la calibration et nécessitent une confirmation de leur composition par les méthodes analytiques de référence.

La distance généralisée selon Mahalanobis (Williams et Norris 1987) est, utilisée pour détecter les spectres aberrants. Ceci est particulièrement utile en prédiction. Ne disposant pas de valeur de référence, il est important de savoir si les échantillons présentés correspondent spectralement à la base ayant servi à l'établissement du modèle.

1.6. Les applications de la SPIR

Des informations complémentaires sur les applications peuvent être obtenues en consultant l'ouvrage de Williams et Norris (1987), qui compte plus de 1.000 références bibliographiques.

La Station de Haute Belgique s'est employée depuis plus de 15 ans à développer des calibrages applicables aux productions agro-industrielles. Par ailleurs, la Station de Haute Belgique a développé des calibrages sur différents types de fourrages (tableaux VIII.3, VIII.4, VIII.5).

Tableau VIII.3 - Calibrages établis sur des herbes de prairie et foin
(régions tempérées)
(Base de données Station de Haute Belgique)

	N	Moy	SD	SEC	R ² _c	SD/SEC	SECV
MAT	315	16.8	6.4	0.73	0.99	8.8	0.83
CBW	308	24.7	5.3	1.10	0.96	4.8	1.19
ADF	248	28.8	5.0	1.32	0.93	3.8	1.47
NDF	252	48.9	7.2	1.61	0.95	4.5	1.78
ADL	251	5.1	1.2	0.57	0.76	2.1	0.61
dMO _{vitro}	264	77.8	8.1	1.81	0.95	4.5	1.95
DMO	337	67.6	6.5	2.0	0.92	3.3	2.15

dMO_{vitro} : digestibilité matière organique in-vitro

DMO : digestibilité matière organique in vivo

Tableau VIII.4 - Calibrages établis sur des fourrages herbacés en
provenance du Togo
(Base de données Station de Haute Belgique)

	N	Moy	SD	SEC	R ² _c	SD/SEC	SECV
MAT	78	6.4	2.4	0.34	0.98	7.1	0.44
CBW	49	31.5	3.9	0.60	0.94	6.5	0.78
ADF	78	40.1	4.8	0.86	0.97	5.6	1.07
NDF	74	67.1	4.7	0.87	0.97	5.4	1.23
ADL	78	5.6	1.6	0.41	0.93	3.9	0.59
dMO _{vitro}	77	38.8	9.9	2.23	0.95	4.4	2.64

Tableau VIII.5 - Calibrages établis sur des feuillages d'arbres méditerranéens (Meuret M. *et al.*, 1993)

	N	Moy	SD	SEC	R ² c	SD/SEC	SECV
MO	204	93.8	3.2	0.54	0.97	5.9	0.63
MAT	206	11.5	5.4	0.69	0.98	7.8	0.88
NDF	90	52.0	13.0	1.36	0.99	9.6	2.06
ADF	110	31.5	10.7	1.85	0.97	5.8	2.36
ADL	108	13.9	6.2	1.04	0.97	6.0	1.48
dMO _{vito}	75	64.0	19.8	1.51	0.99	13.1	2.10

1.7. Abréviations utilisées

Pour l'interprétation des performances des modèles prédictifs, les paramètres suivants sont évalués :

- N : effectif de la population
- Moy : moyenne de la population
- SD : écart-type de la population
- SEC : erreur standard de calibrage
- R²c : coefficient de détermination de calibrage
- SECV : erreur standard de croisée de validation
- R²v : coefficient de détermination
- SD/SEC : rapport écart-type résiduel à l'écart-type de la population.

2. APPLICATION AUX FOURRAGES LIGNEUX TROPICAUX : RESULTATS ET DISCUSSIONS

L'ensemble des calibrages ont été établis sur base des échantillons du projet et des analyses de référence effectuées par les différents partenaires. Les paramètres mesurés par rapport à la teneur en matières azotées (fractions azotées) ont été exprimés en pourcentage de la matière sèche (p.100 MS) (tableau VIII-6).

Après broyage sur moulin Cyclotec (grille 1,0 mm) les échantillons sont mesurés dans la gamme spectrale de 1 100 à 2 500 nm sur un spectromètre NIRSystems 5 000. L'ensemble des données spectrales et analytiques ont été traitées en utilisant le logiciel développé par Infrasoft International (Shenk *et al.*, 1991).

La procédure de calibration est la régression au sens des carrés partiels (Partial Least Squares ou PLS) décrite par Williams et Norris (1987). Le nombre de termes (facteurs) à introduire est déterminé par validation croisée. Cette procédure (Shenk et Westerhaus, 1991) consiste à choisir 2/3 des échantillons sur lesquels un modèle est développé, celui-ci est ensuite appliqué au 1/3 des échantillons restants. La procédure est répétée deux fois en réservant toujours 1/3 des échantillons pour la validation. Les écarts des trois prédictions sont cumulés pour obtenir une erreur standard de validation dont le minimum fixe le nombre de facteurs à introduire dans le modèle. Le modèle final est alors recalculé sur l'ensemble des échantillons.

Tableau VIII.6 - Liste des paramètres dont le transfert sur analyseur proche infrarouge a été étudié

Abréviations	Constituants	unités
MM	matières minérales totales	% m.s.
MO	matières organiques	% m.s.
MAT	matière azotée totale*	% m.s.
MGE	matière grasse (extrait éthéré)	% m.s.
CBS	cellulose brute (scharrer)	% m.s.
CBW	cellulose brute (Weende)	% m.s.
NDF	neutral detergent fiber	% m.s.
ADF	acid detergent fiber	% m.s.
ADL	acid detergent lignin	% m.s.
HEMIC	hémicellulose (NDF - ADF)	% m.s.
CEL	cellulose Van Soest (ADF - ADL)	% m.s.
EB/10	énergie brute	0.1 cal/g MS
TANPT	tannins	
MAS	matières azotées solubles*	% m.s.
MA _{pep}	matières azotées digestibles à la pepsine*	% m.s.
MA _{PRO1}	matières azotées digestibles à la pronase 1 h*	% m.s.
MA _{PRO2}	matières azotées digestibles à la pronase 24 h*	% m.s.
MA _{adf_{ms}}	matières azotées dans l'ADF	% m.s.
SMS	solubilité pepsine cellulase (ms)	% m.s.
SMO	solubilité pepsine cellulase (mo)	% mo
GT	gastest	ml/200mg/ms
dMO _{GT}	digestibilité de la matière organique	% mo
EM _{GT} /10	énergie métabolisable	kcal/100g ms
dMA _{GT}	dégradabilité de l'azote	% ma

*matières azotées = $N_{\text{Kjeldahl}} \times 6.25$

2.1. Etablissement de modèles généraux de prédiction : comparaison entre valeurs estimées et valeurs mesurées

2.1.1. Prédiction des composés organiques majeurs

Les deux premières séries d'échantillons sont à la base des premières calibrations proche infrarouge dont les résultats ont été détaillés dans le rapport de décembre 1991 (tableau VIII.7).

Tableau VIII.7 - Performances des équations pour la prédiction par SPIR des composés organiques de fourrages ligneux (décembre 1991)

% MS	N	Moy	SD	SEC	R ² _c	SD/SEC	SECV	R ² _v
MO	385	90.8	6.0	1.28	0.93	4.7	1.52	0.90
MAT	393	14.3	6.2	1.13	0.97	5.5	1.35	0.95
MGE	341	4.0	2.6	0.69	0.88	3.8	0.82	0.83
CBW	200	24.3	10.6	2.32	0.95	4.6	2.85	0.92
NDF	180	43.2	14.3	3.42	0.94	4.2	5.04	0.86
ADF	187	32.1	12.7	3.12	0.93	4.1	3.72	0.91
ADL	183	11.4	5.9	1.80	0.84	3.3	2.39	0.73
HEMIC	35	11.7	5.4	2.48	0.66	2.2	3.75	0.28
CEL	39	19.8	8.1	2.26	0.91	3.6	3.17	0.83

Avec l'acquisition d'un plus grand nombre d'échantillons, ces équations ont été confirmées et confortées. Les performances de ces équations ont été débattues lors de la réunion de travail de juin 1993 (tableau VIII.8).

Tableau VIII.8 - Performances des équations pour la prédiction par SPIR des composés organiques de fourrages ligneux (juin 1993)

	N	MOY	SD	SEC	R ² _c	SD/SEC	SECV	R ² _v
MO	869	90.5	5.92	1.48	0.94	4.1	1.54	0.93
MAT	871	13.7	5.10	1.14	0.95	4.5	1.21	0.95
MGE	743	3.7	2.17	0.79	0.88	2.9	0.82	0.85
CBW	658	24.8	9.75	2.03	0.95	4.5	2.48	0.93
ENA	760	48.1	9.07	2.59	0.92	3.3	3.11	0.89
NDF	556	45.3	12.90	3.65	0.92	3.3	4.35	0.89
ADF	555	33.7	11.38	3.45	0.90	3.2	4.06	0.87
ADL	556	13.1	6.39	2.62	0.83	2.4	3.07	0.77
HEMIC	560	11.4	5.27	2.97	0.67	1.7	3.58	0.54
CEL	548	20.5	7.35	1.75	0.94	4.1	2.01	0.93
EB/10	26	477.3	15.1	7.70	0.74	2.0	8.72	0.65
TANPT	147	29.0	2.14	10.1	0.79	2.4	10.4	0.37

La prédiction des valeurs manquantes a été effectuée à l'aide de ces modèles. Le rapport de l'erreur standard de calibration à l'écart-type de la population constitue un bon outil pour évaluer les performances des équations. En effet, ce rapport exempt d'unité, permet la comparaison entre les différentes équations et ce, quel que soit le paramètre évalué. Dès que ce rapport est proche de 3.0, voire supérieur à 3.0 la calibration peut être utilisée.

On peut dès lors classer les différents critères en fonction de leur aptitude à être prédit par SPIR:

critères prédictibles:

MM ou MO, MAT, CBW, ENA, NDF, ADF, CEL, MGE

critères difficilement prédictibles

ADL, Hemic

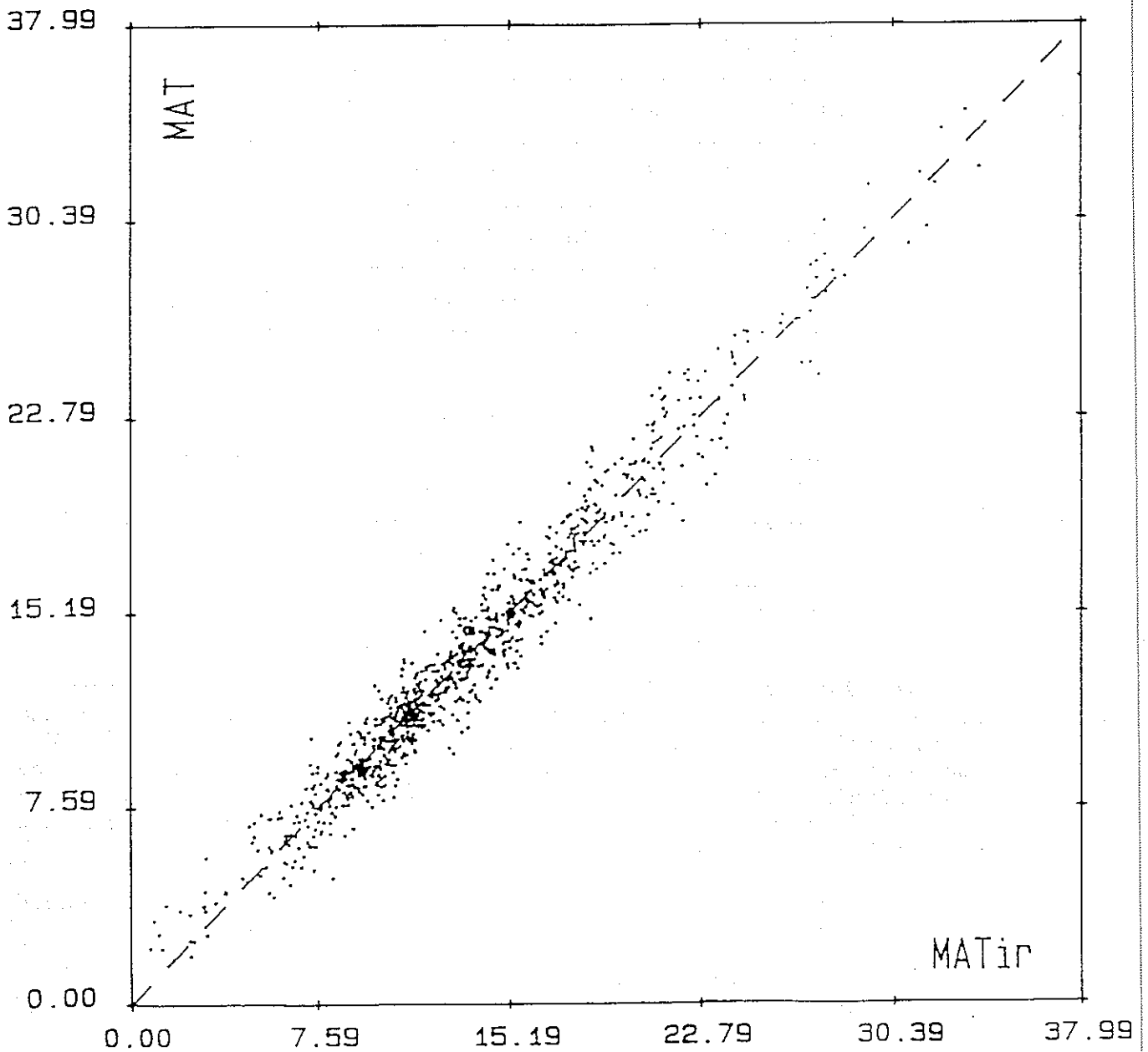
critères nécessitant un complément d'échantillons

EB/10, TANPT

La comparaison entre valeurs prédites et valeurs de référence pour la MAT est reprise à la figure VIII. 2.

Figure VIII.2 - Matières azotées totales : comparaison entre valeurs prédites par SPIR et valeurs de référence

N = 890
SEC = 1.14
 R^2_c = 0.95
SD/SEC = 4.5



De nouveaux essais effectués sur une base de données complémentaires (septembre 1993) ont confirmé ces résultats (tableau VIII.9).

Tableau VIII.9 - Performances des équations pour la prédiction par SPIR des composés organiques de fourrages ligneux (septembre 1993)

	N	Moy	SD	SEC	R ² _c	SD/SEC	SECV	R ² _v
MO	1035	90.5	6.6	1.53	0.92	4.3	1.61	0.91
MAT	1050	13.5	5.6	1.09	0.95	5.2	1.15	0.95
MGE	899	3.8	2.5	0.80	0.86	3.1	0.84	0.84
CBW	808	24.6	9.7	2.32	0.94	4.2	2.48	0.93
CBS	148	20.6	5.0	1.89	0.93	2.6	2.2	0.90
NDF	843	46.1	13.4	4.42	0.89	3.0	4.71	0.87
ADF	827	34.3	12.4	4.12	0.87	3.0	4.43	0.85
ADL	801	13.4	7.5	2.97	0.77	2.5	3.18	0.74
HEMIC	840	10.9	5.8	3.38	0.59	1.7	3.67	0.52
CEL	811	7.7	20.3	1.93	0.93	4.0	2.06	0.92
EB/10	62	465.4	21.7	12.0	0.69	1.8	16.8	0.39
TANPT	296	23.1	21.9	9.2	0.78	2.4	10.2	0.73

La prédiction de la cellulose brute par la méthode Sharrer est plus difficile que la prédiction de la cellulose brute selon la méthode Weende (le rapport SD/SEC étant respectivement de 2.6 et de 4.2). On préférera cette dernière méthode pour l'évaluation de la teneur en cellulose brute⁽¹⁾. Compte tenu du peu de différence entre les performances des dernières équations (septembre 1993) et les précédentes, il ne nous paraît pas utile de refaire les prédictions sur base des derniers modèles. Exception sera faite pour les critères CBS, TANPT et EB.

Le nombre d'échantillons disponibles pour ce dernier critère n'est pas suffisant pour conclure de manière définitive (N=62). Pour ces trois critères de nouvelles prédictions ont été obtenues sur base des modèles développés en septembre 1993.

2.1.2. Prédiction des composés azotés

Lors des premiers essais, les composés ayant trait à la matière azotée étaient exprimés en pourcentage par rapport à la matière azotée totale (tableau VIII.10.). Dans les essais ultérieurs, les teneurs ont été ramenées en pourcentage par rapport à la matière sèche. En effet, la SPIR répond plutôt à une concentration (tableaux VIII. 11 et VIII.12)

Sur base de ces résultats, on peut classer les paramètres selon leur aptitude à être prédit par SPIR.

critères prédictibles

MAS, MApro1_{ms}

critères difficilement prédictibles

MAadf_{ms}

critères nécessitant plus d'échantillons

MApro2_{ms}, MApep

La comparaison entre valeurs prédites et les valeurs de référence de l'azote soluble est reprise à la figure VIII.3.

2.1.3. Prédiction de la digestibilité in-vitro

La digestibilité des fourrages ligneux tropicaux a été approchée sur base de méthodes enzymatiques (SMS, SMO) et sur base de la méthode *in vitro* du "gastest" développée par l'université d'Hohenheim (GT, dMO_{GT}, EM_{GT}/10 et dMA_{GT}). La prédiction de ces critères de digestibilité par SPIR a été étudiée.

(1) la technique de Sharrer est en fait abandonnée au CIRAD-IEMVT depuis 1978.

Tableau VIII.10 - Performances des équations pour la prédiction par SPIR des composés azotés de fourrages ligneux (décembre 1991)

% MAT	N	Moy	SD	SEC	R ² _c	SD/SEC	SECV	R ² _v
S	107	24.1	13.8	4.6	0.84	3.0	6.5	0.78
MAadf _{ma}	159	14.6	8.7	4.1	0.80	2.1	5.2	0.68
MAprol _{ma}	79	36.1	15.6	8.6	0.71	1.8	10.6	0.56
MApro2 _{ma}	43	58.0	13.1	6.0	0.78	2.2	8.8	0.53
MApep _{ma}	61	60.6	21.9	11.9	0.67	1.8	13.9	0.55

Tableau VIII.11 - Performances des équations pour la prédiction par SPIR des composés azotés de fourrages ligneux (juin 1993)

% MS	N	Moy	SD	SEC	R ² _c	SD/SEC	SECV	R ² _v
MAS	125	4.2	3.29	0.82	0.94	3.8	1.21	0.87
MAadf _{ma}	509	2.4	1.47	0.75	0.72	1.9	0.90	0.63
MAprol _{ms}	446	4.3	3.23	0.95	0.90	3.3	1.09	0.88
MApro2 _{ms}	43	9.9	4.45	0.84	0.97	5.0	1.55	0.90
MApep _{ms}	65	10.2	6.51	1.98	0.90	3.5	2.40	0.86

Figure VIII.3 - Azote soluble : comparaison entre valeurs prédites par SPIR et valeurs de référence

N = 125
SEC = 0.82
 R^2_c = 0.94
SD/SEC = 3.8

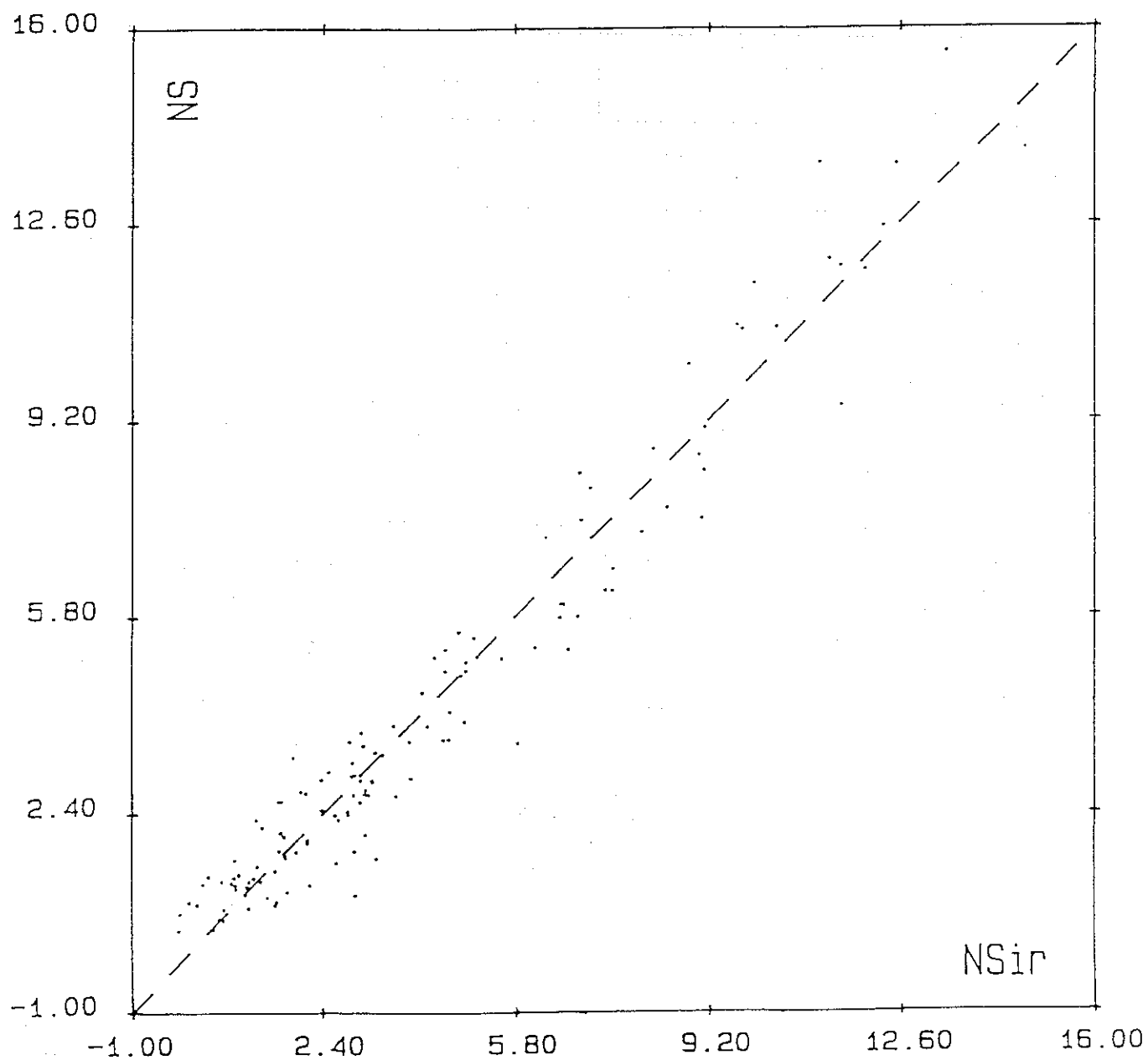


Tableau VIII.12 - Performances des équations de prédiction par SPIR des composés azotés de fourrages ligneux (septembre 1993)

% MS	N	Moy	SD	SEC	R ² _c	SD/SEC	SEV	R ² _v
MAS	129	4.1	3.6	1.02	0.90	3.5	1.27	0.85
MAadf _{ms}	792	2.4	1.8	0.80	0.71	2.3	0.86	0.67
MApro1 _{ms}	691	4.2	3.4	0.94	0.89	3.6	1.02	0.87
MApro2 _{ms}	59	9.4	4.8	1.23	0.92	4.0	1.32	0.91
MApep _{ms}	67	10.2	6.4	1.45	0.95	4.4	2.00	0.90

Les résultats obtenus sont repris aux tableaux VIII.13, VIII.14 et VIII.15. Les premiers résultats obtenus pour les critères SMS et SMO se sont confirmés avec l'augmentation de l'échantillonnage. Ces deux critères sont aisément prédictibles par SPIR. Parmi les critères liés au gastest, la dMO_{GT} est le critère le plus aisé à prédire.

La prédiction des critères GT et, EM_{GT} et dMA_{GT} est plus délicate. La relation entre valeurs prédites et valeurs de référence est reprise aux figures VIII.4 et VIII.5 pour SMO et dMO_{GT}.

Tableau VIII.13 - Performances des équations de prédiction par SPIR de la digestibilité de fourrages ligneux (décembre 1991)

	N	Moy	SD	SEC	R ² _c	SD/SEC	SEV	R ² _v
SMS	147	56.0	15.3	4.10	0.93	3.82	5.29	0.89
SMO	144	53.2	16.0	4.24	0.93	3.8	5.29	0.90

Tableau VIII.14 - Performances des équations pour la prédiction par SPIR de la digestibilité de fourrages ligneux (juin 1993)

	N	Moy	SD	SEC	R ² _c	SD/SEC	SEV	R ² _v
SMS	403	54.5	16.90	4.46	0.94	3.5	5.55	0.89
SMO	400	51.7	17.68	4.80	0.93	3.5	5.87	0.89
GT	451	26.2	11.90	4.11	0.88	2.8	4.87	0.83
dMO _{GT}	455	50.9	13.17	3.79	0.91	3.3	4.36	0.88
EM _{GT} /10	451	167.5	44.39	16.11	0.86	2.7	18.42	0.82

Tableau VIII.15 - Performances des équations pour la prédiction par SPIR de la digestibilité de fourrages ligneux (septembre 1993)

	N	Moy	SD	SEC	R ² _c	SD/SEC	SEV	R ² _v
SMS	597	59.0	17.5	5.22	0.91	3.4	5.71	0.89
SMO	586	51.8	18.3	5.43	0.91	3.4	6.10	0.88
GT	906	23.4	11.4	4.54	0.82	2.5	4.77	0.80
dMO _{GT}	835	48.5	12.4	3.97	0.89	3.1	4.22	0.88
EM _{gt} /10	721	158.6	43.3	16.35	0.84	2.6	17.35	0.82
dMA _{GT}	147	12.8	14.6	6.2	0.77	2.4	7.77	0.65
MAD _{GT} (g/kg ms)	142	19.9	29.2	11.1	0.77	2.0	14.6	0.60

Figure VIII.4 - Solubilité de la matière organique : comparaison entre valeurs prédites par SPIR et valeurs de référence

N = 400
SEC = 4.80
 R^2_c = 0.93
SD/SEC = 3.5

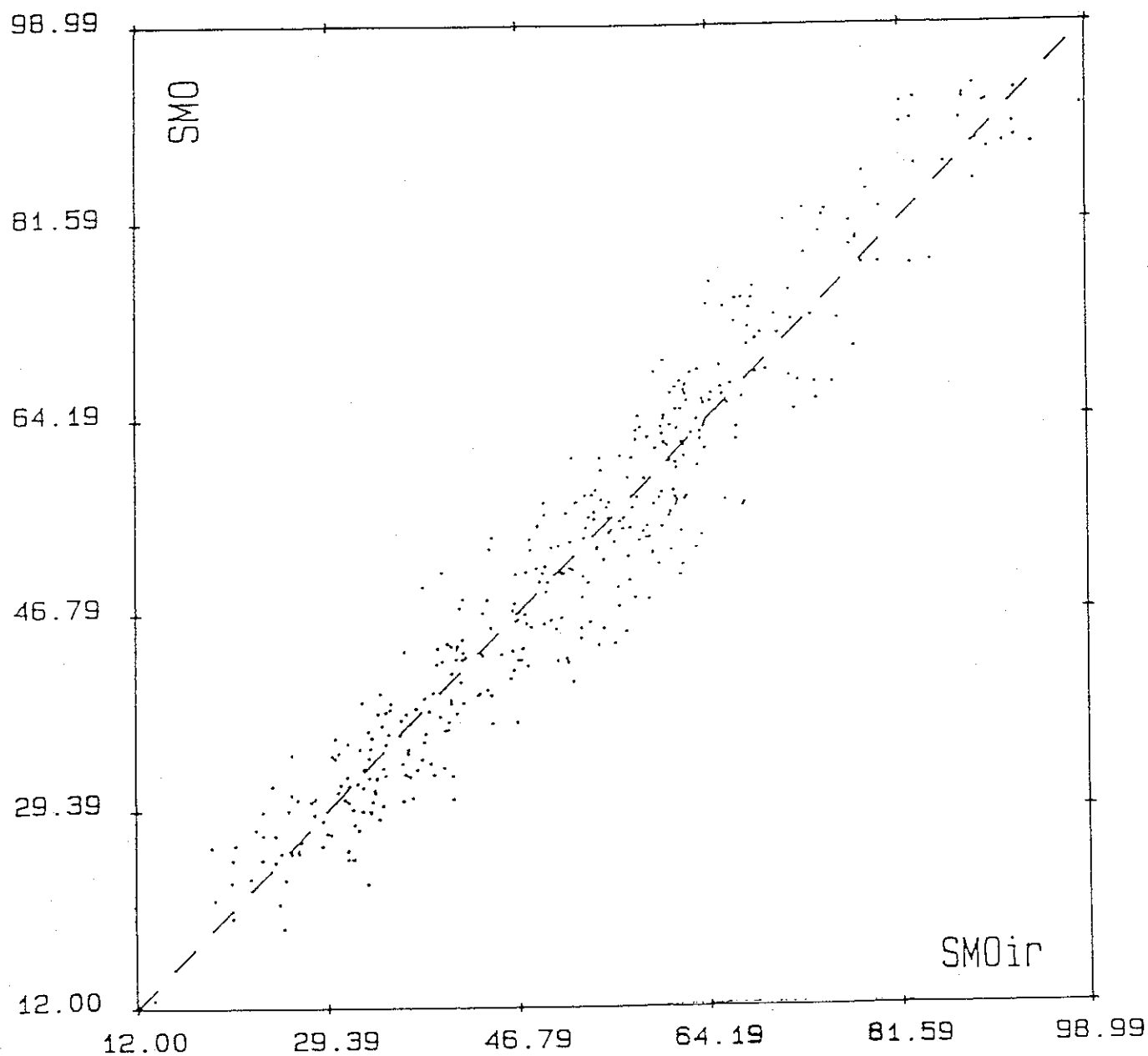
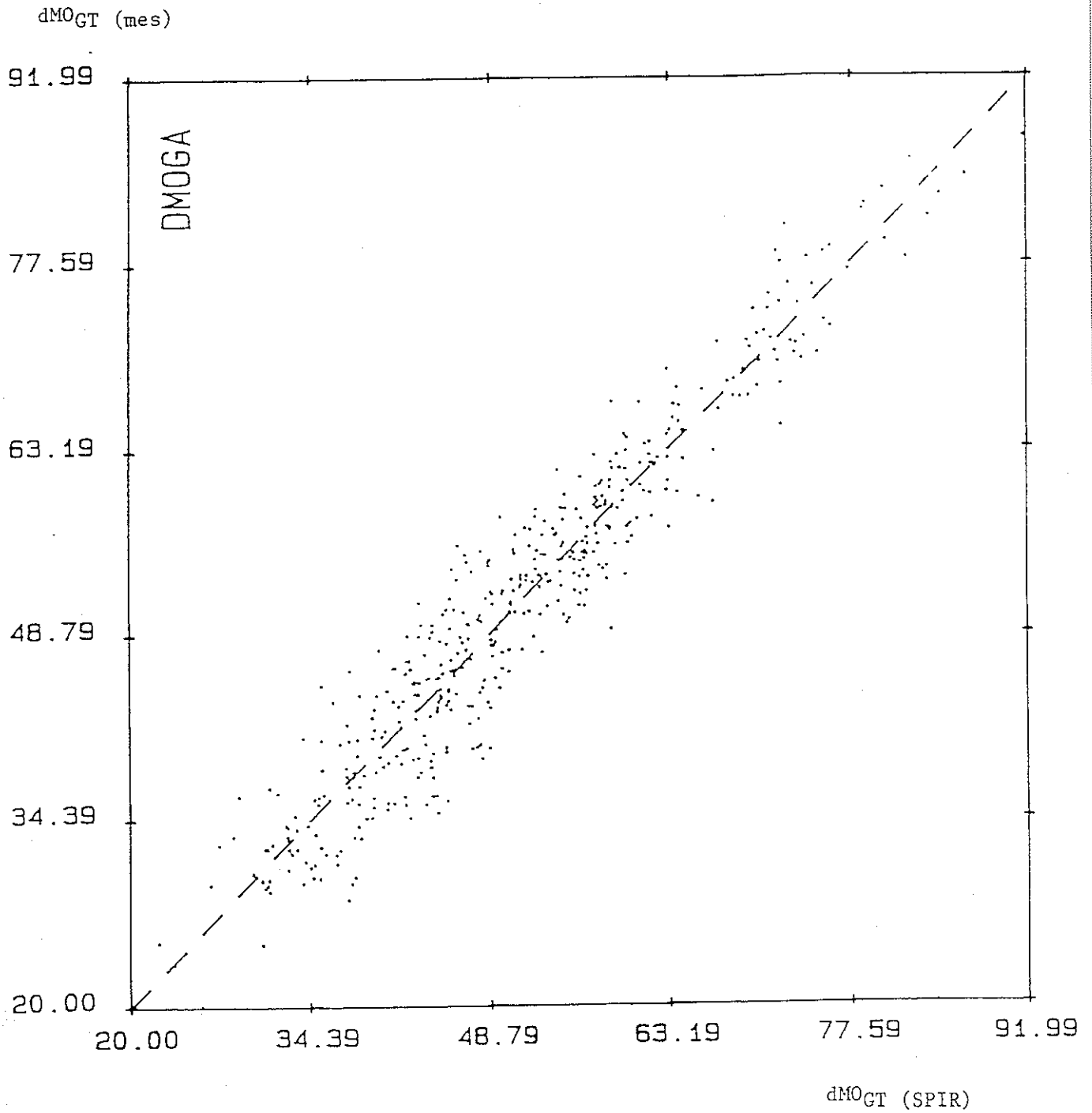


Figure VIII.5 - Digestibilité de la matière organique (gastest) :
comparaison entre valeurs prédites dMO_{GT} (SPIR) par SPIR
et valeurs de référence dMO_{GT} (mes)

N = 455
SEC = 3.80
 R^2_c = 0.91
SD/SEC = 3.3



2.1.4. Prédiction des composés minéraux

La SPIR est essentiellement dévolue à l'analyse des composés organiques (liaisons OH, NH, CH). Les composés minéraux ne répondent pas directement dans la gamme de l'infrarouge proche. Cependant, on peut, dans certains cas, avoir des effets indirects sur des constituants organiques.

Seul le calcium semble pouvoir être dosé par SPIR (tableau VIII.16).

Tableau VIII.16 - Performances des équations pour la prédiction par proche infrarouge des composés minéraux de fourrages ligneux

	N	Moy	SD	SEC	R ² _c	SD/SEC	SEV	R ² _v
K	495	1.2	0.56	0.30	0.71	1.9	0.34	0.64
ICI	411	1.1	1.27	0.72	0.68	1.8	0.87	0.53
Ca	624	1.3	0.96	0.32	0.89	3.0	0.36	0.86
PT	616	0.2	0.07	0.04	0.71	1.8	0.04	0.63
Mg	494	0.3	0.19	0.08	0.82	2.4	0.09	0.77
Cu	127	7.5	3.72	2.55	0.53	1.5	3.36	0.19
Zn	125	25.1	10.85	7.03	0.58	1.5	8.25	0.43
Co	114	0.8	0.93	0.62	0.56	1.5	0.85	0.16
Mn/10	125	14.8	18.48	9.24	0.75	2.0	13.53	0.46
Fe/10	127	59.5	53.34	26.13	0.76	2.0	36.61	0.54
Na/10	70	31.2	56.05	20.21	0.87	2.8	31.75	0.68

2.2. Regroupement d'espèces sur base des données spectrales et analytiques ; établissement d'équations spécifiques

La base de données constituée en cours de projet présente une grande variabilité. Elle intègre un grand nombre d'espèces, de familles, d'organes différents collectés à différents stades de développement, à plusieurs saisons dans plusieurs stations agroclimatologiques et sur un grand nombre d'années. La majorité des échantillons a cependant été réunie de 1991 à 1993.

De ce fait la variation spectrale de l'ensemble des échantillons est importante. L'analyse en composantes principales a montré qu'il fallait 32 axes indépendants pour qualifier cette population. Alors que, bien souvent, 10 à 12 axes suffisent.

Sur base des données spectrales, il n'a pas été possible de constituer des sous-groupes plus homogènes et suffisamment distants l'un de l'autre.

2.2.1. Calibration sur base de groupes

Une procédure "fastclust" du logiciel SAS a permis de distinguer quatre groupes sur base des valeurs analytiques déterminées en laboratoire. Des équations distinctes ont été obtenues pour chacun des groupes. Les performances de ces équations se sont avérées similaires à celles obtenues sur l'ensemble de la population.

La constitution de groupes sur cette base n'a pas permis d'améliorer les résultats pour les critères testés (tableau VIII.17). La distinction de groupes conduit à une réduction de l'effectif ce qui autorise à mettre moins de termes lors de l'établissement des modèles prédictifs (maximum 1 terme pour 10 échantillons).

Tableau VIII.17 - Performances des équations établies après constitution de groupe sur les valeurs laboratoires (Fastelust du logiciel SAS)

p.100 MS	N	Moy	SD	SEC	R ² _c	SD/SEC	SECV	R ² _v
<u>Base totale</u>								
MAT	393	14.3	6.2	1.13	0.97	5.5	1.35	0.95
MGE	341	4.0	2.6	0.69	0.88	3.8	0.82	0.83
NDF	180	43.2	14.3	3.42	0.94	4.2	5.04	0.86
ADF	187	32.1	12.7	3.12	0.93	4.1	3.72	0.91
ADL	183	11.4	5.9	1.80	0.84	3.3	2.39	0.73
<u>Groupe SAS1</u>								
MAT	68	9.1	3.1	1.01	0.87	3.1	1.57	0.69
MGE	68	3.7	1.3	0.78	0.58	2.8	0.90	0.44
NDF	69	37.4	10.0	4.05	0.82	2.5	5.66	0.65
ADF	68	25.0	6.4	2.20	0.88	2.9	3.33	0.73
ADL	71	8.7	2.7	1.31	0.75	2.1	1.75	0.56
<u>Groupe SAS2</u>								
MAT	47	11.0	3.2	0.92	0.90	3.5	1.66	0.68
MGE	51	6.8	3.2	1.52	0.71	2.1	2.09	0.45
NDF	50	37.4	8.0	3.61	0.80	2.2	5.25	0.57
ADF	50	28.7	6.0	2.57	0.81	2.3	3.50	0.65
ADL	50	11.9	3.7	1.54	0.82	2.4	2.25	0.62
<u>Groupe SAS3</u>								
MAT	55	10.3	4.5	1.45	0.90	3.1	1.65	0.87
MGE	55	2.8	2.0	0.48	0.94	4.2	0.77	0.86
NDF	56	59.0	8.9	3.61	0.84	2.5	5.54	0.63
ADF	54	46.7	8.4	2.97	0.87	2.8	4.14	0.75
ADL	54	17.0	6.0	2.65	0.74	2.3	4.01	0.40

2.2.2. Calibrations par famille et par genre

Des essais de calibration de la teneur en protéines ont été effectués en opérant sur la famille des Mimosaceae qui, en nombre d'échantillons, est la plus représentée.

Cette famille a été considérée dans son ensemble en séparant les feuilles des fruits. Le genre *Acacia* a été considéré, séparément en distinguant les feuilles, les fruits et les gousses. Les résultats doivent être pris avec réserve lorsque le nombre d'échantillons est trop réduit ($N < 30$).

Des résultats du tableau VIII.18, il ressort que la distinction entre la famille, le genre, ou les organes ne permet pas d'améliorer la précision des modèles prédictifs.

Tableau VIII.18 - Résultats des calibrations de la teneur en protéines en considérant la famille de *Mimosaceae* et le genre *Acacia*

Population	N	Moy	SD	SEC	R ² _c	SD/SEC	SECV	R ² _v
<u>Base totale</u>	393	14.3	6.2	1.13	0.97	5.5	1.35	0.95
Mimosaceae								
feuilles et fruits	102	16.7	5.2	1.07	0.96	4.9	1.33	0.93
feuilles	39	17.1	5.7	1.56	0.93	3.7	1.79	0.90
fruits	39	15.5	5.1	1.14	0.95	4.5	1.27	0.94
Acacia								
feuilles, fruits, gousses	71	16.4	4.5	0.99	0.95	4.5	1.36	0.91
feuilles	25	14.3	4.8	1.23	0.92	3.9	1.36	0.90
fruits	33	16.5	4.8	1.21	0.94	4.0	1.25	0.93
gousses	19	16.5	5.4	0.84	0.98	6.4	0.87	0.97

2.2.3. Sélections spectrales

Sur base d'une sélection raisonnée de quelques espèces, des groupes ont été constitués en utilisant une sélection basée sur l'analyse en composantes principales des données spectrales.

Les espèces étudiées sont :

Acacia senegal	(AS)
Ficus sycomorus	(FS)
Balanites aegyptia	(BA)
Daniellia oliveri	(DO)
Boscia senegalensis	(BS)
Ziziphus mauritania	(ZM)

A partir de ces six noyaux de départ, six populations ont été constituées sur base de la distance spectrale.

Les distances croisées entre ces différents groupes sont reprises au tableau VIII.19 et constituent la figure VIII.6.

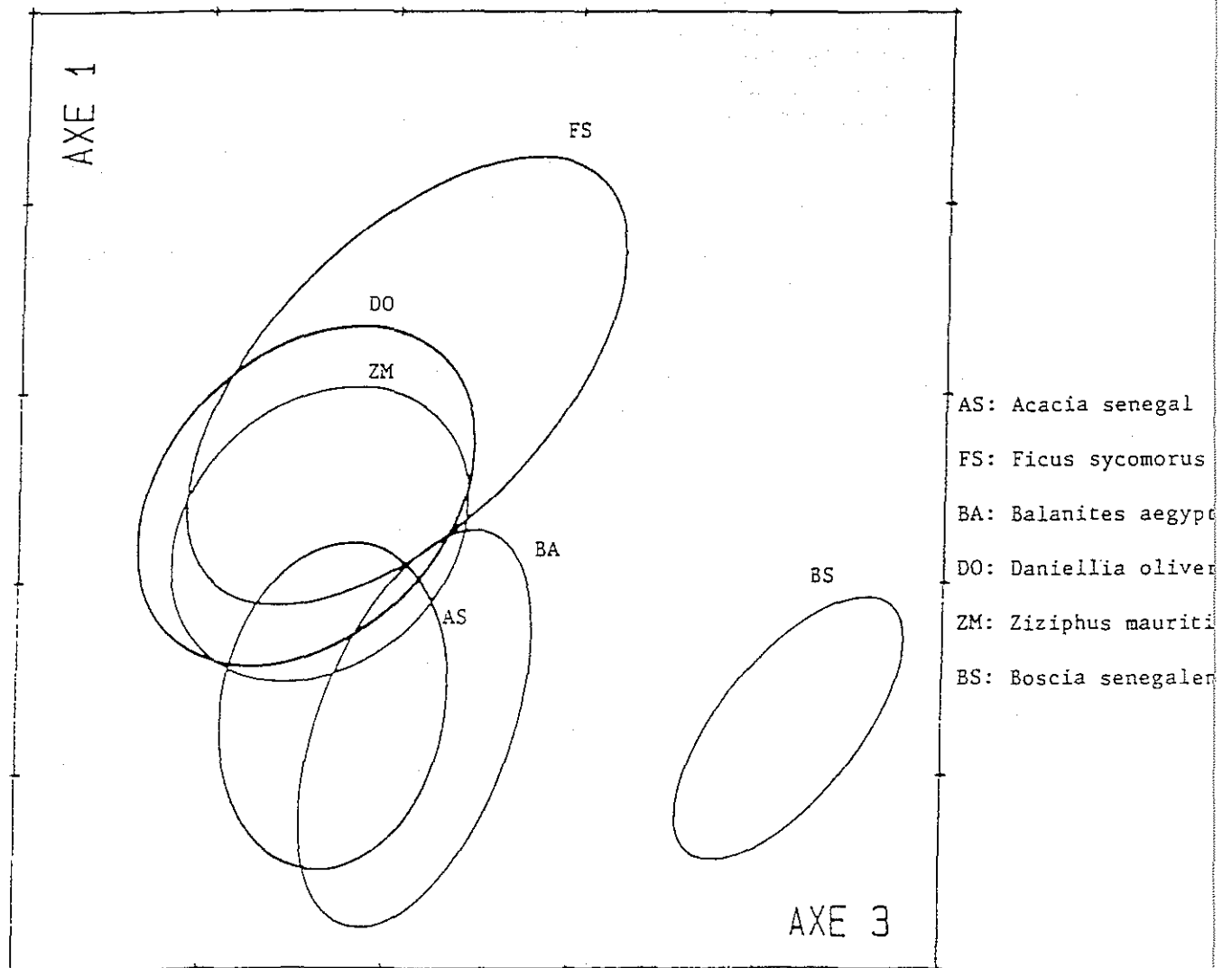
Tableau VIII.19 - Distances généralisées selon Mahalanobis : H croisés

test	réf	AS	FS	BA	DO	BS	ZM
AS		-	3.0	1.9	3.6	12.0	2.2
FS		3.9	-	3.3	2.8	12.7	2.3
BA		3.7	5.6	-	4.2	16.5	4.4
DO		3.0	2.2	2.2	-	8.4	8.1
BS		17.8	11.4	5.2	9.4	-	8.1
ZM		2.3	2.1	1.8	2.1	9.3	-

AS : Acacia senegal
 FS : Ficus sycomorus
 BA : Balanites aegyptia
 DO : Daniellia oliveri
 BS : Boscia senegalensis
 ZM : Ziziphus mauritania

Figure VIII.6 - Visualisation des groupes 1, 2 et 3 selon les axes 1 et 3 de l'analyse en composantes principales sur les spectres obtenus par SPIR

IEMVT : DISTINCTION DE GROUPES



L'étendue des ellipses représente la variabilité au sein du groupe considéré. La variabilité de DO et ZM est pratiquement couverte par FS. Les groupes AS et BA se recouvrent fortement alors que BS est une population peu variable et bien distincte des autres espèces.

Par rapport à celles-ci, BS présente une distance généralisée élevée (tableau VIII.19). Dès lors, trois groupes ont été constitués :

groupe 1: AS, BA
groupe 2: BS
groupe 3: FS, DO, ZM.

Cette démarche a été étendue à d'autres espèces :

Parinari curatellifolia (PA)
Piliostigma thonningii (PI)
Pterocarpus erinaceus (PT)
Salvadora persica (SP)
Spondias mombin (SM)

Sur base des distances généralisées (H croisés) il ressort que le groupe 2 est très distinct des autres et que seule l'espèce SP se distingue des trois autres groupes. Dès lors, on a considéré les trois groupes précédents et le groupe 4 constitué de l'espèce Salvadora persica. Cette espèce est bien distincte et présente une grande plage de variabilité (tableau VIII.20 et figure VIII.7).

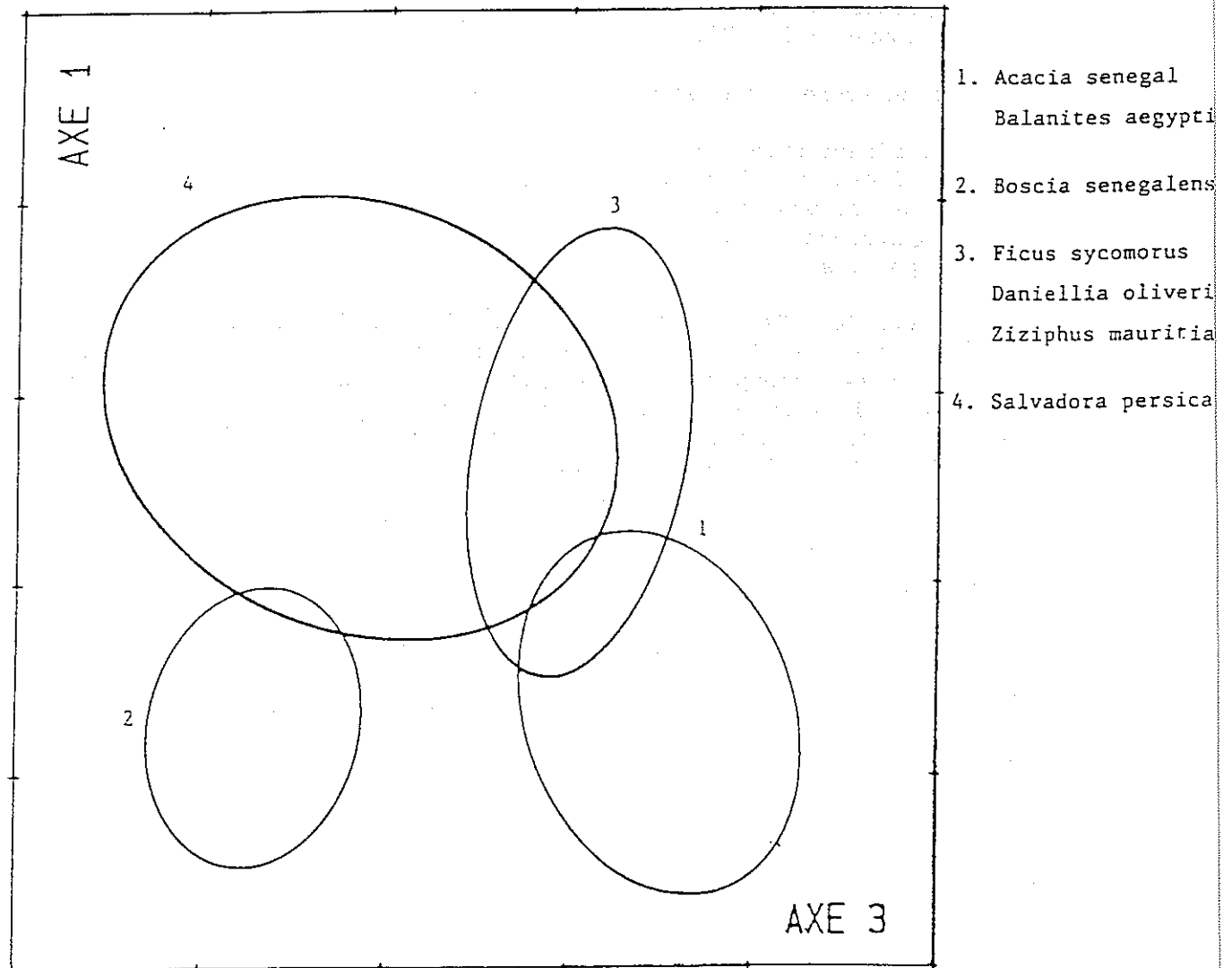
Tableau VIII.20 - Distances généralisées selon Mahalanobis : H croisés

test	réf	GR1	GR2	GR3	PA	PI	PT	SP	SM
GR1		-	12.0	2.8	4.3	2.3	2.4	3.1	3.6
GR2		8.5	-	7.5	7.4	12.7	3.7	2.9	6.1
GR3		2.7	12.2	-	2.6	2.2	2.7	1.7	2.0
PA		2.7	13.2	1.8	-	1.8	2.1	1.7	2.2
PI		2.8	11.5	2.4	3.8	-	2.2	2.4	2.1
PT		2.7	11.3	2.9	4.5	1.9	-	3.9	3.1
SP		9.6	16.5	9.3	10.9	9.5	9.6	-	7.2
SM		3.2	11.8	2.7	3.1	2.2	2.0	1.6	-

ESPECES	GR
AS : Acacia senegal	1
BA : Balanites aegyptia	1
BS : Boscia senegalensis	2
FS : Ficus sycomorus	3
DO : Daniellia oliveri	3
ZM : Ziziphus mauritania	3
PA : Parinari curatellifolia	
PI : Piliostigma thonningii	
PT : Pterocarpus erinaceus	
SP : Salvadora persica	4
SM : Spondias mombin	

Figure VIII.7 - Visualisation des groupes 1 à 4 selon les axes 1 et 3 de l'analyse en composantes principales sur les spectres obtenus par SPIR

IEMVT DISTINCTION DE GROUPES



Sur chacun de ces groupes, des modèles prédictifs ont été établis (tableau VIII.21).

Les effectifs par groupe sont plus réduits et autorisent moins de termes lors de l'établissement du modèle (soit 1 terme par 10 échantillons au maximum).

Dans de rares cas, les précisions sont améliorées : le critère MAT du groupe 3, GT des groupes 2 et 4, dMO_{GT} du groupe 4. Pour les autres cas les précisions sont ou stables ou diminuées. Au vu de ces essais, il ne nous paraît pas judicieux d'effectuer des calibrations par groupe. On peut en rester à des équations générales valables pour une grande diversité d'échantillons.

La distinction des groupes sur base des spectres a été utilisée pour établir des régressions liant la dMO_{GT} aux autres paramètres.

Tableau VIII.21 - Etablissement de modèles prédictifs sur chacun des groupes botaniques obtenus par analogie spectrale

	N	Moy	SD	SEC	R ² _c	SD/SEC	SECV	R ² _v
Population totale								
MAT	871	13.7	5.1	1.14	0.95	4.5	1.21	0.95
GT	451	26.2	11.9	4.11	0.88	2.8	4.87	0.83
dMO_{GT}	455	50.9	13.2	3.79	0.91	3.3	4.36	0.88
$EM_{GT}/10$	451	167.5	44.4	16.11	0.86	2.7	18.42	0.82
Groupe 1								
MAT	109	18.7	6.1	1.47	0.94	4.1	1.63	0.93
GT	96	30.2	6.9	3.28	0.74	2.1	4.03	0.60
dMO_{GT}	87	57.4	6.3	2.67	0.79	2.4	3.30	0.68
$EM_{GT}/10$	87	189.9	27.2	15.27	0.68	1.8	17.61	0.60
Groupe 2								
MAT	37	19.0	3.5	1.06	0.89	3.3	1.20	0.86
GT	36	24.1	3.9	1.26	0.86	3.1	1.64	0.77
dMO_{GT}	35	51.8	5.4	1.53	0.90	3.5	2.13	0.82
$EM_{GT}/10$	33	159.2	19.3	7.23	0.83	2.7	9.60	0.71
Groupe 3								
MAT	216	12.5	3.7	0.72	0.96	5.1	0.83	0.94
GT	181	20.5	8.9	3.30	0.86	2.7	3.86	0.81
dMO_{GT}	167	46.3	9.4	2.91	0.91	3.2	3.35	0.88
$EM_{GT}/10$	153	141.9	32.8	13.40	0.83	2.4	16.12	0.76
Groupe 4								
MAT	40	12.8	4.1	1.14	0.91	3.6	1.74	0.80
GT	35	34.3	9.1	2.50	0.93	3.6	3.30	0.87
dMO_{GT}	35	67.5	14.9	1.81	0.99	8.2	2.55	0.97
$EM_{GT}/10$	28	185.4	28.3	14.53	0.65	2.0	17.29	0.50

2.3. Prédiction de la dégradabilité in vitro par le gastest (dMO_{GT}) à partir des valeurs de composition chimique ou de dégradabilité enzymatique prédites par SPIR

A partir des modèles SPIR établis, les échantillons sont prédits pour chacun des constituants (MO, MAT, MGE, NDF, ADF, ADL, HEMIC, CEL, MAS, MA_{ADFms} , MA_{PRO1ms} , MA_{PRO2ms} , EB, SMS, SMO, TANPT, CBW, MA_{pep}) (VI.6.2.c). A partir de ces valeurs prédites, nous avons cherché à estimer la dMO_{GT} par des régressions multilinéaires. Celles-ci ont été menées en considérant l'ensemble des échantillons (groupes 1, 2, 3 et 4) et en considérant les groupes séparément (tableau VIII.22.).

groupe 1 : *Acacia senegal* (AS) *Balanites aegyptia* (BA)
 groupe 2 : *Boscia senegalensis* (BS)
 groupe 3 : *Ficus sycomorus* (FS)
 Daniellia oliveri (DO)
 Ziziphus mauritiana (ZM)
 groupe 4 : *Salvadora persica* (SP)

Faisant abstraction de la dispersion (SD) de chacune des populations, on améliore la précision de la relation en considérant les groupes séparés. Lorsque l'on intègre la dispersion de chacun des groupes, l'intérêt n'est plus si net. Compte tenu de ces résultats, il ne nous paraît pas indispensable d'opérer de distinction entre les différents groupes.

Lors de l'établissement de ces régressions, le critère SMO ou SMS intervient à chaque fois EB et MAS ressortent dans trois équations, TANPT et MA_{PRO2} interviennent dans deux équations.

Sur l'ensemble des groupes la dMO_{GT} peut être estimée à partir des valeurs prédites de SMS, EB, MAS, TANPT.

Tableau VIII.22 - Régressions multilinéaires en vue d'estimer la dMO_{GT} sur base des valeurs prédites des autres paramètres

groupes	N	Moy	SD	ETR	R^2	SD/ETR	Critères
1 à 4	591	50,5	10,9	3,21	0,91	3,4	SMS, EB, NSms, TANPT
1	135	57,2	5,9	2,90	0,74	3,0	SMO, MA_{PRO1} , MA_{PRO2} , MA_{pcp}
2	43	52,1	4,7	1,26	0,93	3,7	EB, MAS, SMO
3	368	45,8	8,8	2,66	0,91	3,3	EB, SMS, MAS, TANPT
4	50	65,6	15,0	3,13	0,96	4,8	MA_{pcp} , SMS, MA_{PRO2} , MO

2.4. Essais sur une espèce importante : *Faidherbia albida*

Faidherbia albida constitue une espèce qui revêt un intérêt particulier dans la zone géographique étudiée. Cette espèce étant représentée dans la base de données (76 échantillons dont 40 de feuilles, 32 de fruits et 4 échantillons d'organes indéfinis), son étude plus approfondie a été menée. Les résultats des prédictions obtenues sur des modèles généraux établis sur l'ensemble des données disponibles (tableau VIII.23) ont été comparés avec des modèles établis sur la seule espèce envisagée (tableau VIII.24). Il ressort de ces essais que les prédictions obtenues sur modèles spécifiques conduisent à de meilleurs résultats. Pour autant qu'une espèce soit suffisamment représentée et pour autant qu'elle présente une variabilité spectrale et chimique suffisante, on peut arriver à améliorer les précisions en établissant des modèles spécifiques. Ceci n'est pas nécessairement généralisable à toutes les espèces étudiées dans le cadre du projet.

2.5. Essais d'un mode de calibration particulier la "Local Weighted Regression"

Avec l'accroissement des performances des micro-ordinateurs, de nouveaux concepts de calibration ont vu le jour. L'un de ceux-ci, la "Local Weighted Regression" ou calibration locale est particulièrement adapté à de grosses bases de données présentant une grande variabilité spectrale. Avec près d'un millier d'échantillons de différentes espèces végétales, la base de données disponibles dans ce projet rencontre ces caractéristiques.

Ce mode de calibration repose sur le principe suivant :

1. On dispose d'un set restreint d'échantillon à prédire (quelques dizaines d'échantillons).

2. On dispose d'une large base de données spectrales et analytiques (de l'ordre du millier).

3. Par analogie spectrale, on recherche, pour chaque échantillon à prédire, les n échantillons dans la base de données qui lui sont les plus proches (de l'ordre de la centaine d'échantillons à optimiser).

4. On établit un modèle prédictif sur ces n échantillons et on prédit l'échantillon inconnu. Par cette technique, chaque échantillon est en prédiction pure et a une calibration qui lui est spécifique. En outre, cette technique met mieux en évidence des échantillons présentant une grande discordance entre les valeurs laboratoires et les valeurs prédites.

Dans le cadre de cette étude, les aberrants spectraux ont été écartés lorsque la distance généralisée de Mahalanobis excède 3.0. Les aberrants laboratoires ont été écartés en prédisant toute la base de données en mode local. Une centaine d'échantillons ont été sélectionnés pour constituer un set de prédiction pure et le reste sert à la calibration. Les performances des prédictions en mode local sont comparées à celles obtenues par voie classique. Le tableau VIII.25 montre que dans la majorité des cas, les performances obtenues en mode local sont supérieures. Compte tenu de la lenteur de ce mode de calibration, nous avons limité les possibilités au niveau du traitement mathématique des spectres (correction de ligne de base et dérivée première) et du nombre d'échantillons choisi ($n = 150$).

Tableau VIII.23 - Prédiction des échantillons de *Faidherbia albida* sur modèles existants établis sur l'ensemble de la base de données

	N	R ²	SEP	Pente	Biais
MO	75	0.91	1.15	0.99	0.49
MAT	75	0.90	0.95	0.99	-0.43
MGE	75	0.89	1.37	1.40	0.74
NDF	45	0.85	3.12	0.86	1.21
ADF	74	0.66	5.09	0.59	0.13
ADL	74	0.55	3.48	0.89	1.91
HEMIC	45	0.04	2.96	0.24	-1.12
CEL	74	0.88	2.15	1.08	0.82
SMS	44	0.67	6.37	0.79	0.97
SMO	44	0.61	7.22	0.79	0.79

Tableau VIII.24 - Précisions des modèles établis sur les échantillons de *Faidherbia albida*

	N	SEC	R ² c	SECV	R ² v
MO	75	0.51	0.98	0.87	0.94
MAT	71	0.53	0.93	0.84	0.81
MGE	72	0.38	0.98	0.45	0.97
NDF	45	2.00	0.92	3.45	0.77
ADF	68	1.20	0.92	1.86	0.80
ADL	73	1.61	0.79	2.04	0.66
HEMIC	45	2.02	0.27	2.45	0.00
CEL	74	0.94	0.97	1.55	0.93
SMS	43	1.77	0.97	3.53	0.86
SMO	41	1.97	0.93	2.83	0.86

Tableau VIII.25 - Comparaison entre des modèles établis en mode local et par régressions multilinéaires classiques (PLS)

	MODE LOCAL			MODE PLS		
	SEP	R ²	SD/SEP	SEP	R ²	SD/SEP
MO	1.25	0.95	4.3	1.77	0.90	3.0
MAT	1.25	0.94	4.1	1.24	0.94	4.1
MGE	0.65	0.92	5.9	1.05	0.80	3.6
NDF	3.83	0.91	3.4	5.43	0.84	2.4
ADF	3.29	0.92	3.5	4.32	0.88	2.7
ADL	2.31	0.91	3.3	3.23	0.84	2.4
HEMIC	2.98	0.65	1.7	3.44	0.52	1.4
CEL	1.87	0.90	3.1	2.19	0.86	2.7
GT	3.58	0.90	3.1	4.86	0.81	2.3
DMOgt	3.26	0.93	3.8	4.72	0.85	2.6
EMgt/10	14.26	0.88	2.8	18.72	0.79	2.2

3. ESSAIS SUR ANIMAUX : DEGRADABILITE *IN SACCO*, DIGESTIBILITE *IN VIVO*

Sur base de données analytiques obtenues sur animaux les paramètres suivants ont été testés :

DTMA : dégradabilité théorique de la MAT, *in sacco*
 DTMS : dégradabilité théorique de la MS, *in sacco*
 DMSV : digestibilité de la MS, *in vivo*
 DMOV : digestibilité de la MO, *in vivo*

Compte tenu du faible effectif disponible (N=7), il n'a pas été possible d'établir un modèle prédictif pour le critère DTMS. Pour les autres critères, les effectifs sont réduits et les modèles obtenus doivent être considérés comme orientatifs et susceptibles d'être améliorés en augmentant l'effectif (tableau VIII.23). Une augmentation d'effectif permet de mettre plus de termes dans l'équation de prédiction et obtenir ainsi une équation plus robuste.

Il semblerait au vu de ces résultats que le critère DMOV soit le plus aisément prédictible.

Pour comparaison, la prédiction de la digestibilité de la matière organique *in vivo* d'herbe et de foin a conduit aux résultats suivants:

N = 337, SEC = 2.00, $R^2c = 0.92$, SD/SEC = 3.3.

Compte tenu du faible échantillonnage disponible et de la grande diversité de la base de données à prédire, il nous paraît hasardeux d'utiliser ces modèles en prédiction. Ces modèles demandent à être confirmés par un plus grand nombre d'échantillons représentatifs de la population des fourrages ligneux étudiée.

Tableau VIII.26 - Prédiction de la digestibilité des fourrages ligneux (*in sacco* et *in vivo*)

	N	Moy	SD	SEC	R^2c	SD/SE C	SECV	R^2v
DTMA	19	58,0	18,9	12,0	0,60	1,6	13,9	0,43
DMSV	33	49,1	10,7	5,7	0,71	1,9	6,2	0,66
DMOV	30	53,6	11,6	4,2	0,82	2,7	4,7	0,76

CONCLUSION

En général, les corrélations entre les valeurs obtenues par méthodes de référence et prédites par SPIR sont bonnes pour la plupart des composés organiques. Cependant, comparés à des résultats antérieurs (tableaux VIII.3, VIII.4 et VIII.5), les résultats que nous obtenons sont moins précis (tableau VIII.12 et VIII.15) pour les paramètres MAT, SMO (ou DMO_{gt}), MO et moins encore pour des paramètres tels que NDF, ADF, ADL.

Une partie de l'explication réside dans la grande hétérogénéité botanique des échantillons et celle des organes prélevés. De plus, ils ont été prélevés sur un grand nombre d'années ce qui constitue une source supplémentaire d'hétérogénéité.

Des divergences dans le protocole de récolte et de conditionnement des échantillons peuvent également interférer. Le manque de reproductibilité des méthodes de référence à travers les années peut également influencer les performances des équations.

La constitution de groupes ou de sous-groupes en fonction de la nature des échantillons, de leur composition ou sur base d'une classification spectrale ne suffisent pas à améliorer les niveaux de précision. Il est nécessaire, pour cela, de travailler au niveau d'une espèce végétale (cas de *Faidherbia albida*) ou d'utiliser des techniques particulières de calibration (Local Weighted Regression) particulièrement adaptées à de grosses bases de données présentant une grande variabilité botanique et, de ce fait, spectrale.

Il n'en reste pas moins, que compte tenu, de la plage de variation enregistrée pour ce type de fourrage, la SPIR est une bonne méthode de prédiction des constituants organiques et des digestibilités *in vitro* ; cette méthode peut donc contribuer à la connaissance de la valeur nutritive des réserves fourragères ligneuses en vue d'une meilleure gestion de celles-ci.

La digestibilité de la matière organique par le gastest peut être prédite directement à partir des données spectrales ($SEC = 3.79$, $R^2_c = 0.91$, $SD/SEC = 3.3$).

Une autre approche consiste à prédire SMS, EB, MAS, TANPT et à estimer la dMO_{gt} par régression multilinéaire (tableau VIII.22).

REFERENCES BIBLIOGRAPHIQUES

- BERTRAND D., LILA M., FURTOSS V., ROBERT P., DOWNEY G. (1987) - Application of principal component analysis to the prediction of lucerne forage protein content and in vivo dry matter digestibility by NIR spectroscopy. *Journal of the Science of food and agriculture*, 41, 299-307.
- BISTON R., DARDENNE P. (1984) - Prediction of forage digestibility with near infrared reflectance spectroscopy (NIRS). *Proceedings CEC meeting*, Brussels 4-5 sept.
- BISTON R., DARDENNE P. (1985) - Application de la spectrométrie de réflexion dans le proche infrarouge. Prévision de la qualité des fourrages en vue de leur exploitation rationnelle. *Bulletin Recherche Agronomique, Gembloux* 20 (1/2), 23-41.
- BISTON R., DARDENNE P. (1990) - Contribution de la Spectrométrie dans l'infrarouge proche pour l'analyse de la qualité des productions agricoles. *Analisis*, 18 (10), 27-30.
- DARDENNE P., BISTON R. (1986) - Attempt to recognize wheat species by discriminant analysis. *International NIR/NIT Spectroscopy Conference*, Budapest-Hungary, 12-16 May.
- DARDENNE P. (1990) - Contribution à l'utilisation de la spectrométrie dans le proche infrarouge pour l'étude de critères de qualité des céréales et des fourrages. Thèse de doctorat, Faculté des Sciences Agronomiques B-5030 Gembloux - Belgium.
- DARDENNE P., SINNAEVE G., BISTON R., LECOMTE P. (1991) - Evaluation of NIT for predicting fresh forage quality. *4th International NIR/NIT Spectroscopy Conference Aberdeen-Scotland*.
- DARDENNE P., BISTON R., SINNAEVE G. (1992) - Calibration transferability across NIR instruments. *5th International NIR/NIT Spectroscopy conference*, Haugesund-Norway, 16-20 June.
- HELLEMANS P., DARDENNE P., RINES A., DIENG A., LOBET P., HAINEAUX E., BISTON R., COMPERE R. (1990) - Prédiction de la composition chimique des fourrages tropicaux par la spectrométrie de réflexion dans le proche infrarouge. *Bulletin Recherche Agronomique, Gembloux*, 25 (4), 473-488.
- MEURET M., DARDENNE P., BISTON R., POTY O. (1993) - The use of NIR in predicting nutritive value of MediterRanean tree and shrub foliage. *Journal of Near Infrared Spectroscopy*, 1, 45-54.

SHENK J.S., WESTERHAUS M.O., ISI-NIRS-2 (1991) - Software for Near-Infrared Instruments. Infracore International, User Manual.

SHENK J.S., WESTERHAUS M.O. (1991) - Crop Science 31 : 1548.

SINNAEVE G., DARDENNE P., BISTON R. (1993) - Prediction of antinutritional factors in pea by Near Infrared Reflectance Spectroscopy (NIRS). 2nd ANF's workshop, Wageningen - The Netherlands, 1-3 December.

WILLIAMS P., NORRIS K. (1987) - Near Infrared technology in the agricultural and food industries. American Association of cereal chemists, inc., 330 p.

ANNEXE

SPECTROMETRIE DE REFLEXION DANS LE PROCHE INFRAROUGE

RESUME PAR H. GUERIN ET P. LEFEVRE

La méthode consiste à analyser le spectre de radiation diffuse d'une source de lumière monochromatique de longueur d'onde comprise entre 1 100 et 2 500 nm réfléchi par les particules du fourrage.

Les données spectrales de fourrages de composition connue (cf. chapitres V à VII) servent à la mise au point de modèles de calibration² par le calcul des régressions multiples.

Les modèles de calibration sont en général établis à partir d'une centaine d'échantillons, 2/3 servant à bâtir le modèle, le 1/3 restant à le valider. Ces groupes sont alors permutés deux fois (cross-validation). Les écarts entre les données analytiques et les valeurs prédites sont exprimés en terme d'écart-type ou erreur standard³.

La précision des modèles est déterminée par les coefficients de détermination, de calibration et de validation ainsi que par les erreurs standard qui intègrent aussi les erreurs de laboratoire.

Dans un premier temps, l'ensemble des données spectrales analytiques est utilisé pour la construction des modèles puis le fichier est épuré sur deux critères :

- élimination des échantillons spectralement éloignés du centre de gravité de la population. Ils sont identifiés par analyse en composantes principales et calcul de la distance généralisée de Mahalanobis (Williams et Norris, 1987). Cette distance est, pour un spectre donné, sa distance en unité d'absorbance par rapport à la moyenne de tous les spectres pour une ou plusieurs longueurs d'onde données. Ces distances mesurées en terme de "H" doivent autant que possible rester inférieures à 3 (Station de Haute Belgique, 1991) ;
- élimination des échantillons dont l'écart entre les valeurs analytiques et les valeurs prédites est supérieur à deux fois l'erreur standard de validation (SEV).

² A l'aide du logiciel Infracsoft (Shenk, 1989).

³ - SEC = erreur standard de calibration (précision théorique du modèle).
- SECV = erreur standard de cross validation (précision du modèle en cross validation).
- SEP = erreur standard de prédiction (sur un set d'échantillons totalement indépendants).

Il est aussi possible, en particulier dans le cas d'une population botaniquement très hétérogène, comme celle des fourrages ligneux étudiée dans le cadre du projet ST2/215, de procéder à une classification des échantillons suivant leurs spectres ou sur la base des données des laboratoires par analyse discriminante. Dans les deux cas, les procédures de calibration sont appliquées à chacune des classes considérées individuellement.

La constitution des classes peut aussi reposer sur la description botanique des échantillons.

Pour un groupe botanique donné, il est habituellement nécessaire de disposer d'une centaine d'échantillons représentatifs des variations de composition de la population pour obtenir des calibrations assez précises.

Le fichier du projet ST2/215 malgré le nombre d'échantillons proche de 1 500 n'a pas permis de constituer des groupes spectralement homogènes et permettant des calibrations plus précises qu'avec l'ensemble des données. Pour chaque critère analytique on dispose donc d'un modèle unique de prévision spectrale appliqué à l'ensemble des espèces (tableaux VIII.9, 12 et 15).

Enfin, dans le cas particulier du fichier de fourrages ligneux du projet ST2/215 compte tenu des difficultés à obtenir des modèles de prédiction suffisamment précis avec l'ensemble du fichier, même divisé en classes, une procédure de calibration locale (Local Weight Regression) a été appliquée. Cette procédure consiste à rechercher pour chaque échantillon dont on veut prédire les caractéristiques, un sous-ensemble d'échantillons du fichier de calibration dont les caractéristiques spectrales sont les plus proches. Cette procédure a globalement permis une amélioration de la prévision des prédictions (tableau VIII.26).

Dans la pratique de l'utilisation des résultats de calibration et de prédiction, il est possible pour un groupe botanique donné de corriger les prévisions spectrales par des équations de régression entre les données spectrales et les données mesurées au laboratoire (tableau VIII.27 relatif à *Faidherbia albida*, par exemple).

Ces équations sont plus ou moins précises suivant les espèces végétales et les critères analytiques.

L'examen des paramètres statistiques attachés à ces régressions permet de juger de l'intérêt, pour une espèce particulière, des données spectrales ou au contraire de la nécessité de procéder à l'analyse au laboratoire :

- un coefficient de détermination très hautement significatif, une pente proche de 1, une ordonnée à l'origine proche de 0 indiquent que les prédictions de CB par le modèle général peuvent être utilisées sans erreur importante ;

- la dMO_{GT} des gousses peut être prédite (avec précision : $etr = 1.5$ points) à partir de la prédiction SPIR (modèle général : $etr = 4$ points) moyennant une correction ;

$$dMO_{GT \text{ lab}} = 0.80 \ dMO_{GT \text{ SPIR}} + 11.3$$

ce n'est pas le cas pour la dMO_{GT} des feuilles, dont la prédiction n'est pas améliorée par l'équation de correction : la précision reste médiocre (ETR

Tableau VIII.27 : Composition chimique et dégradabilité enzymatique ou *in vitro* (gas test) des feuilles et gousses de *Faidherbia albida* échantillonnées dans le cadre du projet ST2/215 et analysées en spectrométrie infrarouge : relations entre les valeurs mesurées au laboratoire et celles prédites en fonction de modèles de calibration établis pour l'ensemble des espèces ligneuses du projet.

En p.100 de la Matière sèche sauf * % Matière Azotée ** % Matière Organique	FEUILLES									GOUSSES								
	Echantillons étudiés en NIR Résultats laboratoire					Prédiction NIR Xlab = a XNIR + b				Echantillons étudiés en NIR Résultats laboratoire					Prédiction NIR Xlab = a XNIR + b			
	Moy.	E.T.	Min.	Max.	N	a	b	ETR	R2	Moy.	E.T.	Min.	Max.	N	a	b	ETR	R2
ANALYSES ET DÉGRADABILITÉ AU LABORATOIRE																		
<u>Matière Organique</u>	89	2	85	92	38	0,69	28,0	0,9	0,72	96	1	94	96	30	0,29	68,5	0,4	0,26
<u>Constituants pariétaux :</u>																		
.Cellulose Brute	21	3	16	24	8	1,1		2,1	0,99	28	2	25	32	10	1,02		1,5	0,99
.Neutral Detergent Fiber	39	4	31	49	28	0,69	13,2	2,0	0,75	43	3	37	48	10	0,44	24,2	2,4	0,47
.Acid Detergent Fiber	29	4	20	39	38	0,64	12,0	2,5	0,58	34	4	26	42	29	0,42	18,3	2,5	0,57
.Acid Detergent Lignin	14	2	9	22	38	0,57	8,1	1,4	0,65	9	2	6	12	29	0,37	5,4	1,2	0,45
.ADL/ADF %	49	3	40	55	38					26	3	21	32	29				
<u>Matières Azotées :</u>																		
.Totales	13	2	10	17	38	1,11	-2,0	0,5	0,92	10	2	6	15	30	1,17	-2,1	1,0	0,68
.Dégradables par la Pronase	3	1	2	4	38	-	-	-	NS	7	2	2	10	28	1,08	0,6	1,4	0,42
.Dégradables par la Pronase *	25	4	14	32	38	-	-	-	NS	62	13	41	99	28	1,21		11,5	0,97
.Liées à l'ADF	2,7	0,8	1,6	5,2	38	0,58	8,9	0,7	0,32	1,1	0,3	0,7	1,7	29	-	-	-	NS
.Liées à l'ADF *	22	5	10	34	38	0,67	8,3	3,6	0,51	11	3	6	19	29	-	-	-	NS
<u>Tannin précipitant la</u> <u>seroalbumine bovine</u>	3,1	1,0	1,6	4,5	8	0,12		0,6	0,97	0,7	0,4	0,2	1,3	10	-	-	-	NS
<u>Matière Grasse</u>	6	1,5	4	10	38	1,08	1,2	0,8	0,67	1,2	0,3	0,7	2,5	30	-	-	-	NS
<u>Dégradabilité de</u> <u>la Matière Organique</u>																		
.par la Pepsine Cellulase - SMO **	52	6	34	60	22	0,58	19,6	2,5	0,81	63	6	51	76	16	0,51	33,5	4,2	0,49
.production de gaz	23	6	15	31	8	0,75	6,7	4,0	0,65	39	5	29	48	9	0,56	16,1	1,7	0,90
.estimé <i>in vitro</i> par le gastest - dMOgt **	49	5	42	55	8	-	-	-	NS	57	4	49	66	9	0,80	11,3	1,5	0,90

de 4 à 5 points). En revanche, la prédiction de la SMO des feuilles de *Faidherbia albida* bénéficie de l'application d'une équation de correction : l'ETR passe de 6 à 2.5 points ;

- certains paramètres ne peuvent être prédits dans le cas de *Faidherbia* (coefficients de détermination non significatifs) : en particulier, teneur en tannins précipitants des gousses, dégradabilité enzymatique de l'azote des feuilles, fraction de l'azote des gousses liée aux parois, etc.. Dans ce cas, le recours à l'analyse est indispensable.

L'analyse spectrale, sur la base de la distance généralisée de Mahalanobis et de l'étude des résidus issus de la calibration, permet aussi d'identifier des valeurs d'analyses aberrantes (erreurs de transcriptions ou de résultats ou d'analyses) ou des échantillons atypiques mal identifiés sur le plan botanique. Des corrections ont ainsi pu être faites dans le fichier du projet ST2/215.